Justification for CDF data handling specifications

1. Number of Data Handling Job starts per day –

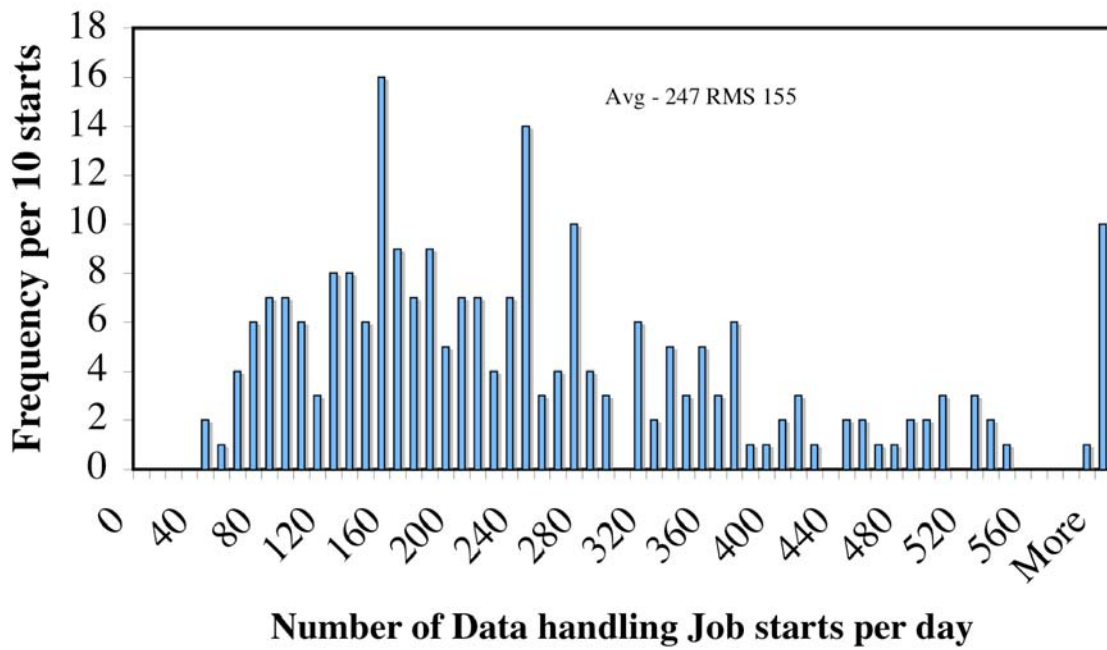**CAF/Condor Caf DH job starts**
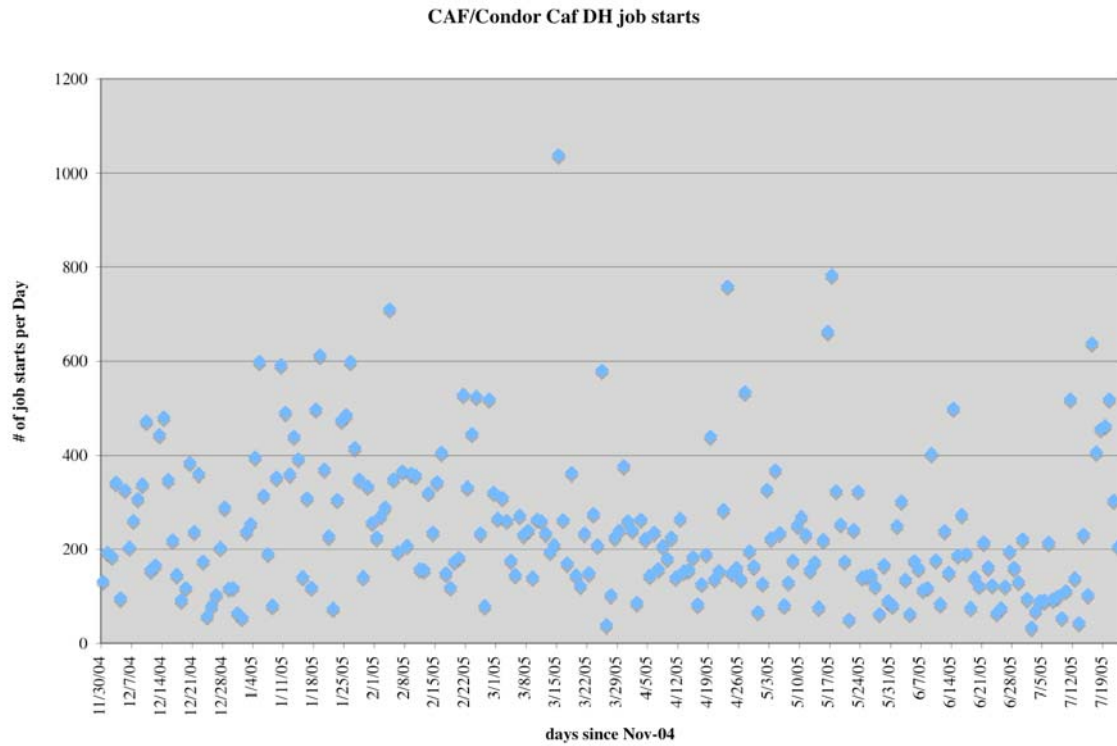
We see from this plot that CDF has a peak number of data handling jobs start of ~1000 Job starts/day.
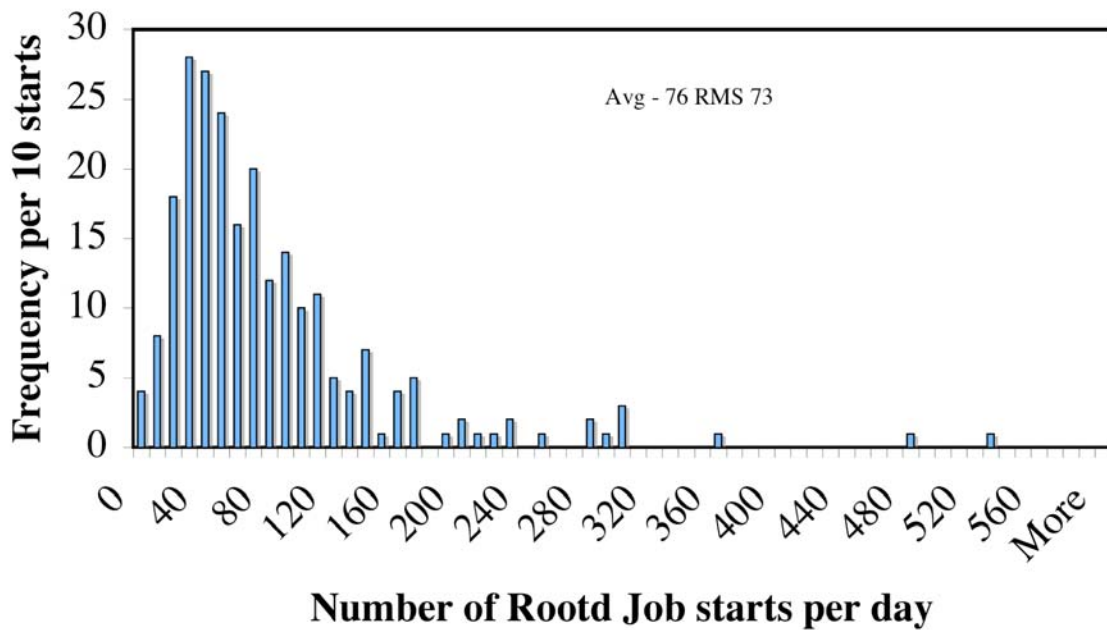
CAF/Condor Caf All job starts
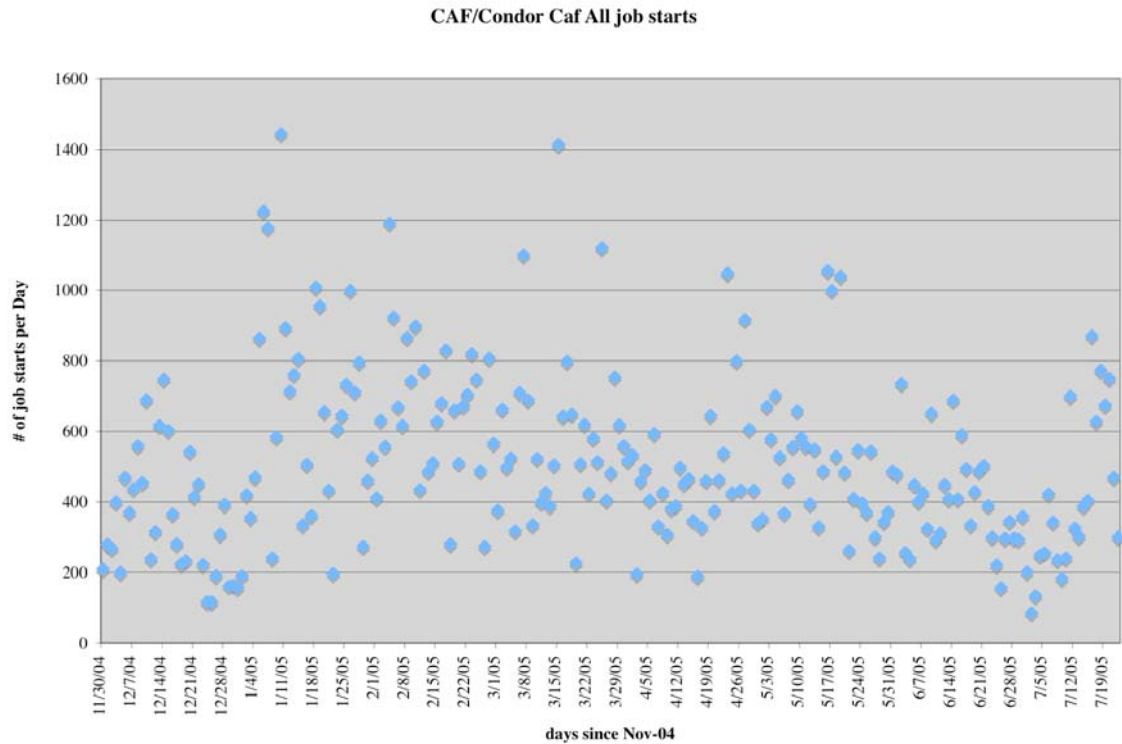
We see from this plot that CDF has had peak number of job starts for jobs of all
Types (MC generation/production, Data processing and RootD analysis) of
1400 job starts per day. This corresponds to ~1 job start per minute.

The distribution of job starts per day for Data Access jobs, MC production/generation and
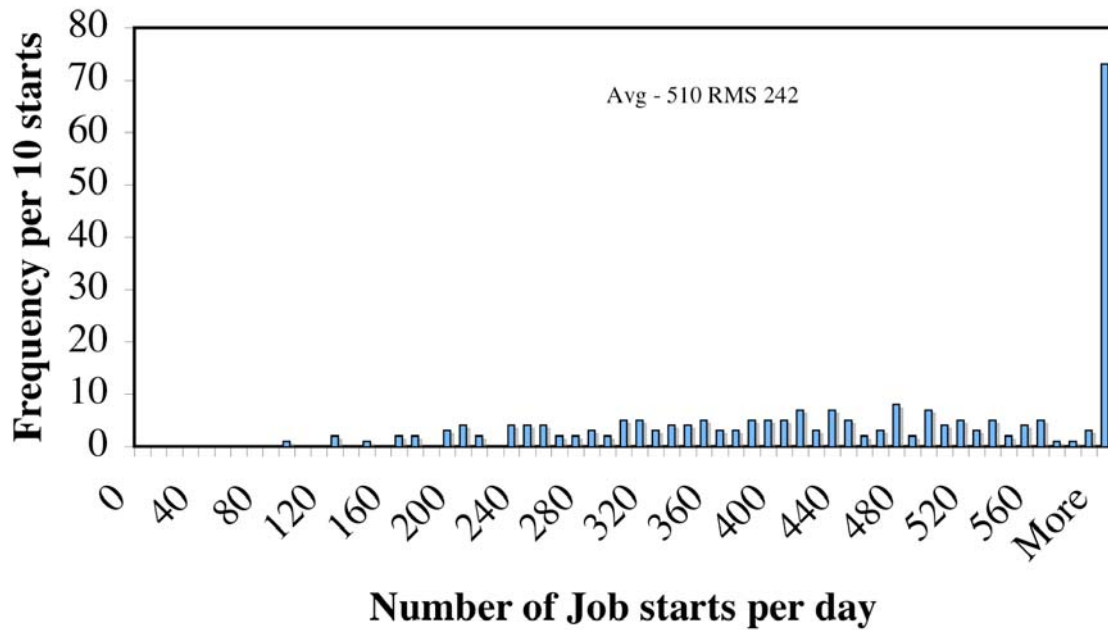rootd jobs is shown in the next three figures

Data Handling Job starts CAF and CondorCAF

Avg - 247 RMS 155

Frequency per 10 starts (y-axis, 0 to 18)

Number of Data handling Job starts per day (x-axis, 0 to More)

RootD Job starts CAF and CondorCAF

Avg - 76 RMS 73

Frequency per 10 starts (y-axis, 0 to 30)

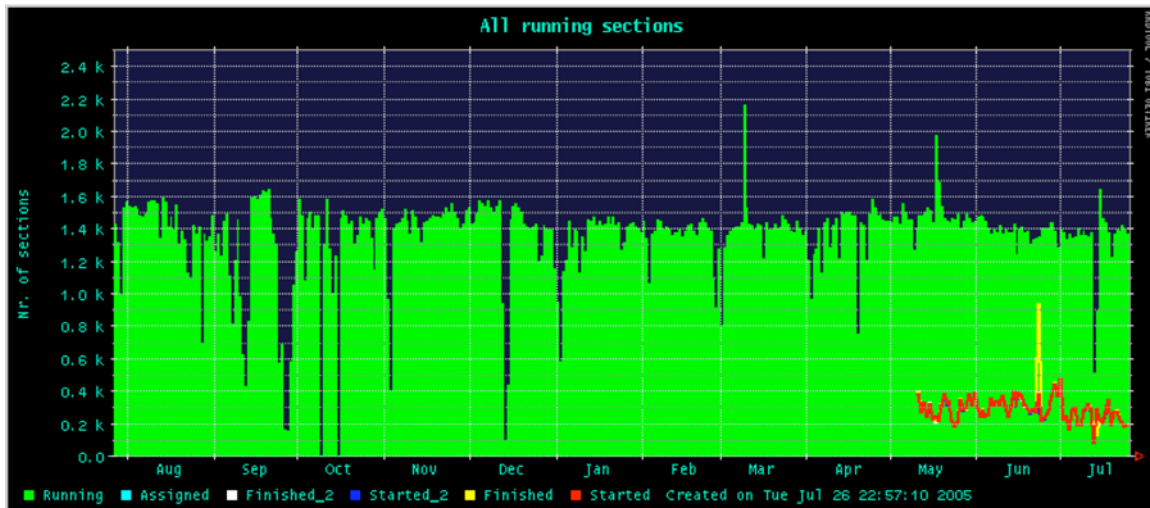**Number of Rootd Job starts per day** (x-axis, 0 to More)
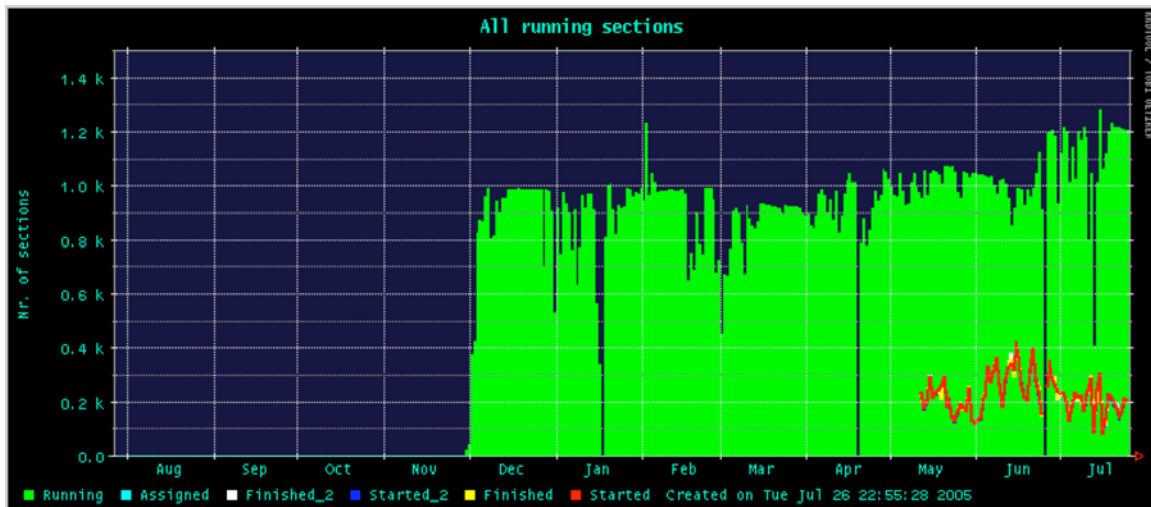
All Job starts CAF and CondorCAF

The number of running job sections for the CAF and Condor CAF is shown in the next two plots.

CondorCaf running sections (Avg. ~ 1400 sections always running)

Caf running sections (Avg. ~ sections always running)

Looking at these two plots I would estimate the number of running sections to
Be ~ 2600 running sections. Approximately 50% of the sections are running jobs that
require the CDF DH resources. This means that on average there are ~ 1300 clients
Running per day. In the near future we expect to group the farm by another 300 nodes
(1200 VM's). Currently there are 2700 VM's between the CAF's.

The number of open files per day can be seen in the read transfers
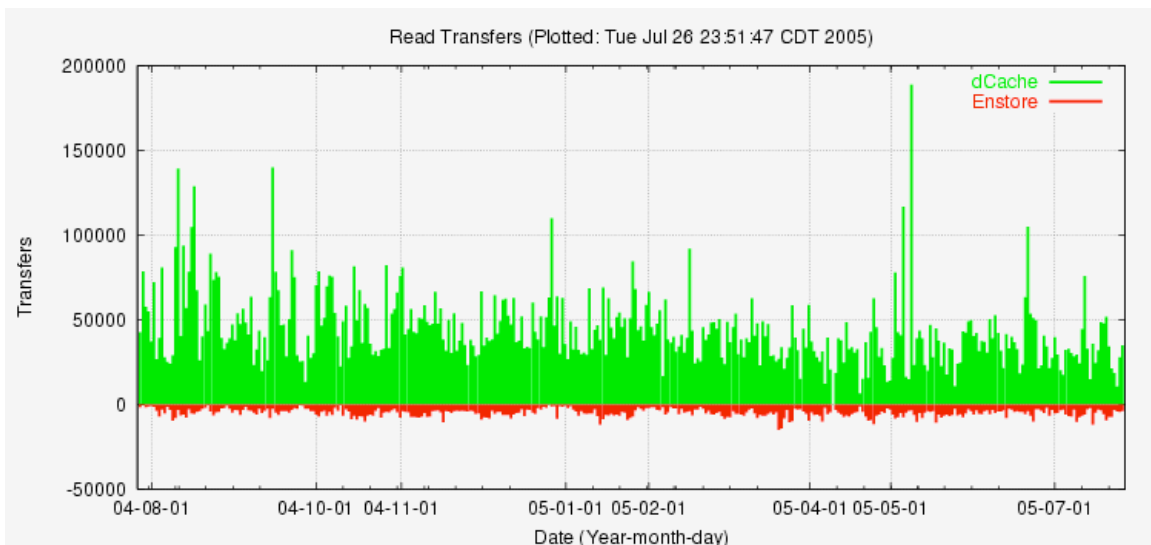
CDF can reach a rate of 180,000 read transfers per day.

From the plot of the Bytes read per day, one can see the CDF reads up to 40 TB when
there is much activity.

Bytes Read (Plotted: Tue Jul 26 23:51:47 CDT 2005)
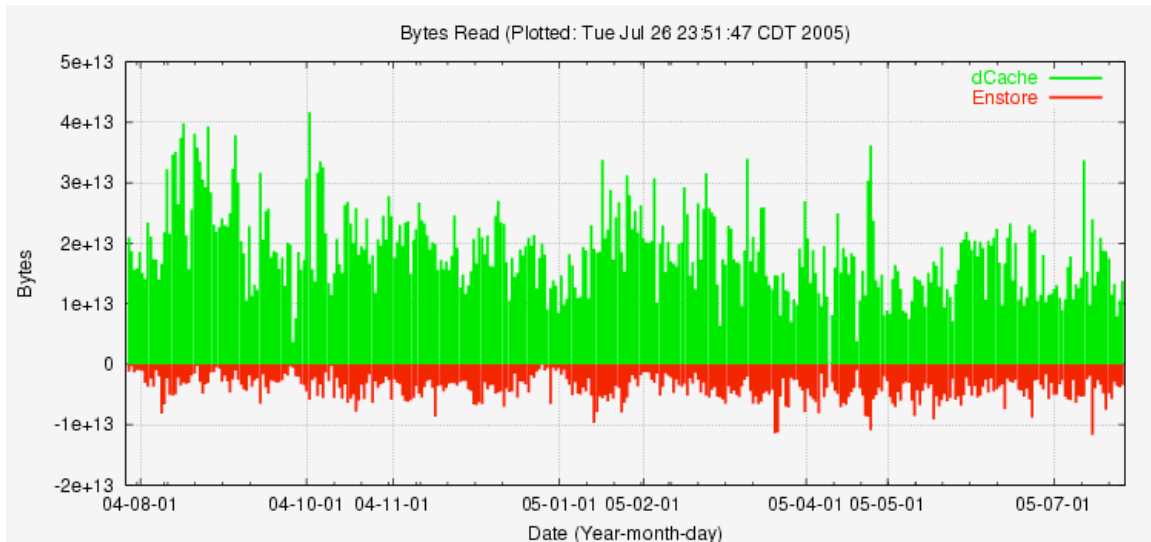
Current Caf/Condor CAF	Daily Rate	2700 VMs rate/minute
Read Transfers	180000	125
Read Volume (TB)	40	0.03
# of DH jobs	1000	0.69
# of DH sections	1300	0.90

CondorCaf/ Caf after upgrade   (1200 addition VM's) (assume all will be used for datahandling)

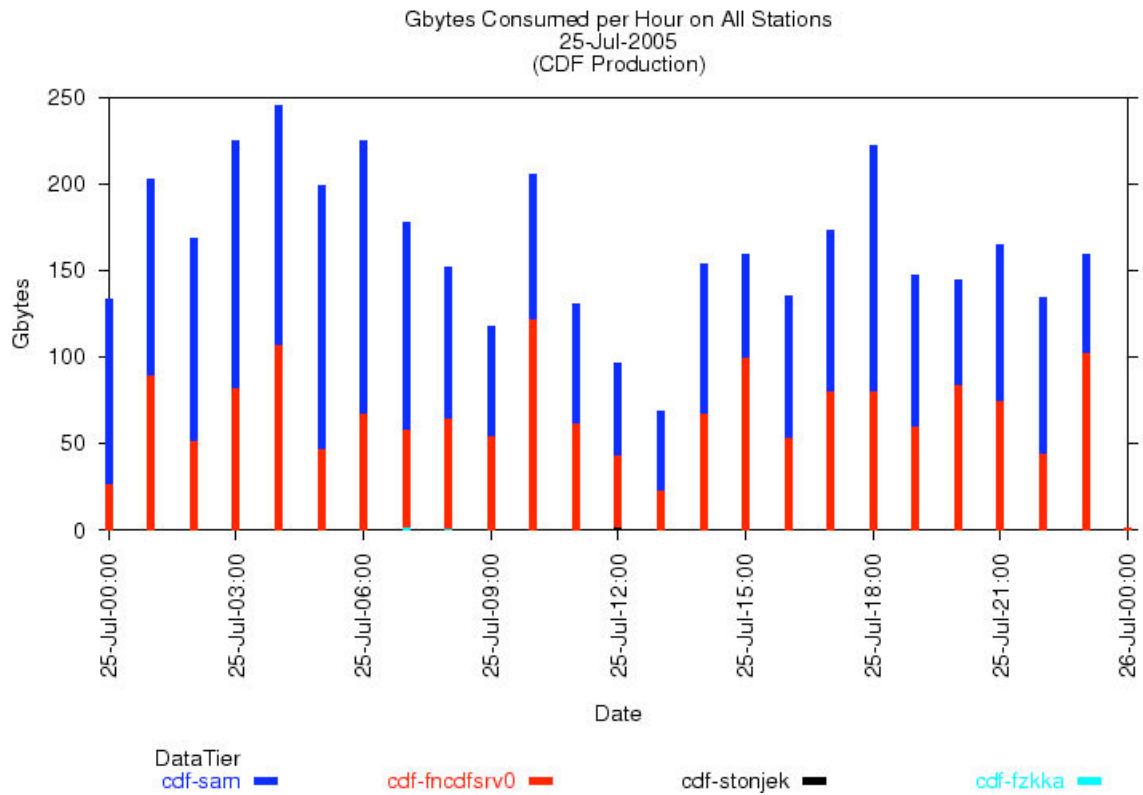|		
Read Transfers	260000	180.56
Read Volume (TB)	57.78	0.04
# of jobs	1444.44	1.00
# of DH sections	1877.78	1.30

# of jobs -> # of project starts/stops
# of DH sections -> # of clients
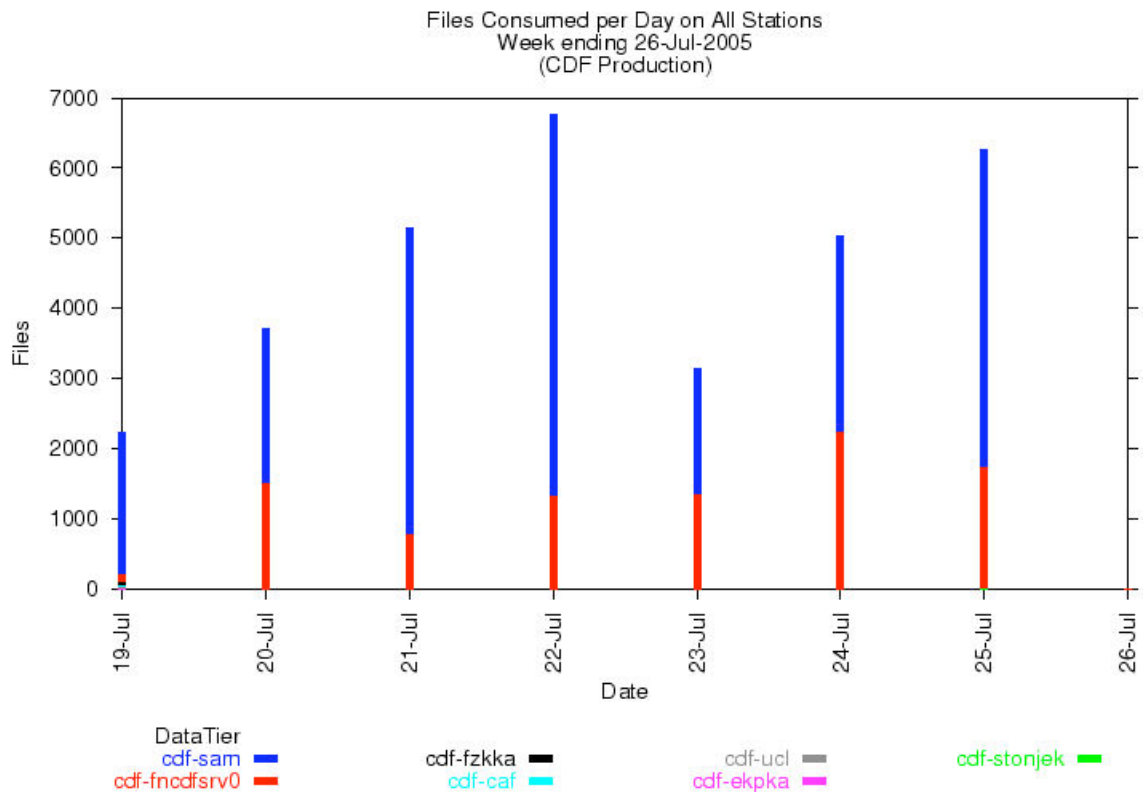Read Transfers -> # of file opens

Production Farm Numbers

For previous week - GBytes consumed (Peak 2 TB/day)

Gbytes Consumed per Day on All Stations
Week ending 26-Jul-2005
(CDF Production)
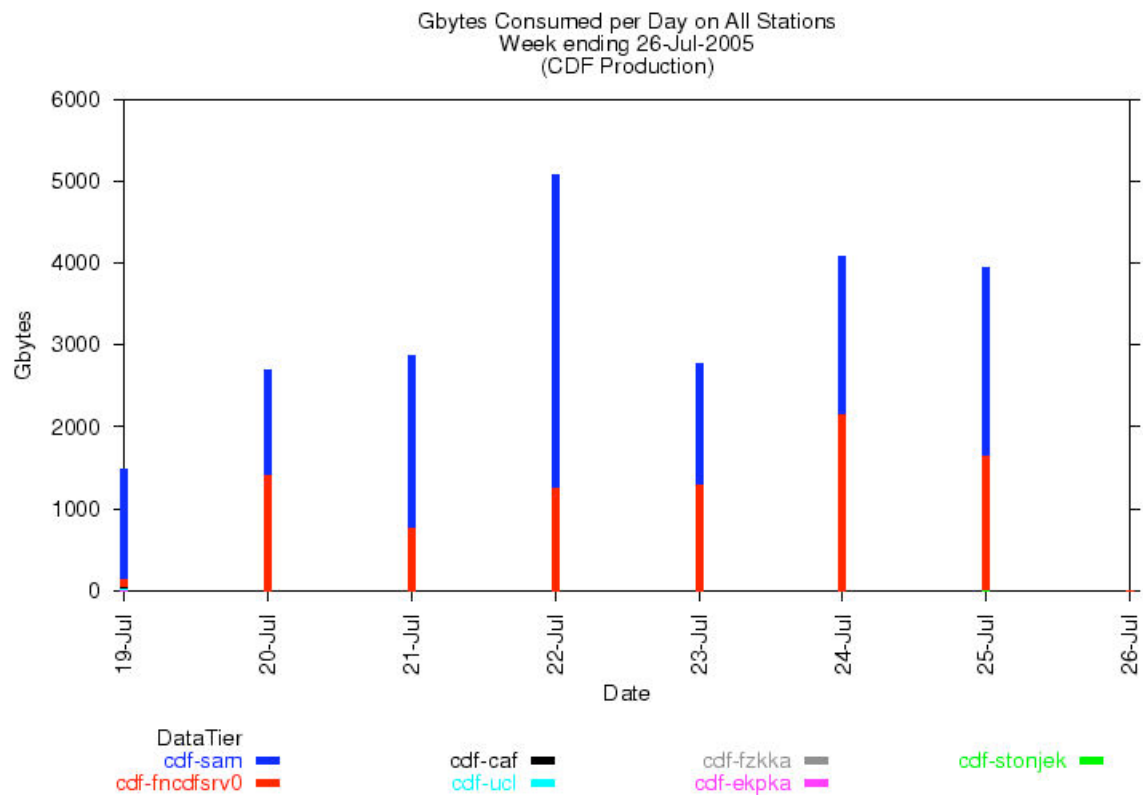
Consumption for past hour

Gbytes Consumed per Hour on All Stations
25-Jul-2005
(CDF Production)

# of files used in previous week:

Files Consumed per Day on All Stations
Week ending 26-Jul-2005
(CDF Production)

# of files Consumed in previous day:

Files Consumed per Hour on All Stations
25-Jul-2005
(CDF Production)